

## Bridging the gap between collaborative and realist evaluation: A general critique and case study of European Union employment policies

Carpenter, Mick

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

Rainer Hampp Verlag

### Empfohlene Zitierung / Suggested Citation:

Carpenter, M. (2005). Bridging the gap between collaborative and realist evaluation: A general critique and case study of European Union employment policies. *International Journal of Action Research*, 1(3), 311-338. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-414162>

### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

# **Bridging the gap between collaborative and realist evaluation: A general critique and case study of European Union employment policies**

Mick Carpenter

After first establishing the need for a reflexive approach to evaluation, the article focuses on the theoretical breakthrough of the 'fourth generation' evaluation paradigm proposed by Guba and Lincoln (1989). This is seen as having much to offer, but their stakeholder focused approach has the danger of lapsing into relativism. Therefore a 'realist' approach is proposed building on Pawson and Tilley (1997), going beyond deconstruction of stakeholder perspectives to make verifiable statements about programme and project effects. This is then worked through with an outline review of the issues associated with evaluating European Union Employment Strategies. The article concludes by suggesting that a realist concern for critical praxis in evaluation can also bring it closer to action research paradigms.

**Keywords:** Evaluation, realism, employment, European Union

## **1. Introduction**

The purpose of this paper is sympathetically to analyse emerging critical traditions in evaluation and find ways of bridging collaborative and 'realist' traditions, working through the concrete example of European employment programmes. I argue that while evaluation should always start with the perceptions and understandings of stakeholders, the evaluator necessarily has an independent role in assessing which have greater plausibility rather than simply

stop at synthesising different accounts. This is because there is the need to allow for the possibility of contradictory or irreconcilable meanings, and also because 'deeper' causal processes may lie at a level not easily accessible to local actors. Thus while evaluation rightly has a pragmatic concern with improving programmes and/or organisations, and helping stakeholders listen and respond to each other, it should not abandon a fundamental concern with uncovering causal processes. 'Realist' approaches, which incorporate social meanings and processes into the causal analysis of change, can help to provide a 'middle way' between extremes of social constructionism and positivism. While the elitist tendencies of objectivist evaluations might be acknowledged, and a preference for participative evaluation from below asserted, there is no need to throw out the scientific baby with the positivist bathwater. In other words, while I sympathise with collaborative evaluation's efforts to bridge gaps between internal actors and external judges, evaluators bring something to different to the table and need to be more than non-judgemental project conciliators.

These reflections arise from evaluative research funded from the European Social Fund (ESF) EQUAL programme undertaken as part of the Great Britain wide SEQUAL project aimed at identifying good practice in integrating excluded and discriminated groups into the labour market. It also arises from discussions held within the evaluation sub-group of SEQUAL's transnational partners in FACETS, which sought to develop cross-national learning about labour market integration of migrants and ethnic minorities (for information on SEQUAL and FACETS see <http://www.surrey.ac.uk/politics/cse/sequal.htm>). The first part of the article seeks to address general issues, and these are then applied in outline to the specific context of the European Union's employment policies.

## **2. Deconstruction and beyond**

At a time when a pragmatic and rather mechanical concern with 'what works' has increasingly come to dominate public policy, seeking to utilise evaluation as a set of depoliticised technologies for evidence based public policy (see for example, Davies/Nutley/Smith 2000), it is necessary to assert explicitly the

political and critical nature of evaluation as a reflexive activity (Taylor/Balloch 2005). The advantage of the constructivist tradition to evaluation is that it brings the political nature of evaluation to the fore and leads to a critical questioning of the way that ideological purposes are incorporated within evaluation strategies.

Thus if evaluation is to be a critical and publicly useful exercise, contributing to democracy and something more than simply a set of commoditised services provided to paying customers, its first task is not to take the objectives and methods of programmes for granted but to interrogate or 'deconstruct' the policy discourses which underpin them. These include sets of values and ideals which at the very least need to be brought into view and scrutinised. This is in contrast with much standard evaluation practice, which is to assume the aims of programmes as given, seeing the evaluator's task as focusing on whether delivery mechanisms are effective. This is often allied to a scientific paradigm of positivism which derives from a notion of the traditional experimental model as the 'gold standard'. The aim is to see whether, controlling for all other variables, there is a measured change that can be assumed to be the result of a 'treatment' or intervention. This calls upon evaluators to be external, expert and impartial judges, and places a premium on their ability to precisely measure the changes from which causal processes can then be inferred.

The reflexive approach by contrast calls on the evaluator to spell out his or her own values, and at the same time to bring w the central value assumptions underpinning programmes into view. It asserts this against the 'value free' tradition which has increasingly been allied to a pragmatic 'Third Way' politics which asserts that the relative success or failure of policies should be empirically validated rather than determined on the basis of either left or right wing ideologies. Policy should be based instead on finding out 'what works' through experimentation and evaluation, and then 'mainstreaming' it. An immediate problem is that there is an assumed consensus that every reasonable person is assumed to agree that crime must be reduced, health improved, economic growth fostered, unemployment reduced, social inclusion fostered, and so on. This is instead of acknowledging that policy action derives from sets of values to which we *choose* to subscribe. This is reminiscent of

Lukes's (1974) 'radical' view of organisational power, which defines it not just in terms of making or preventing things happening but in setting the overall 'taken for granted' and scarcely visible framework in which only certain things are contested within defined bounds.

This is not to suggest that such orthodox evaluation is never challenging, for example, following Cochrane (1972) it 'rigorously' seeks to test out whether medical or other professional pretensions are justified by the practical evidence. It therefore does not assume that anyone has prior knowledge, without it being demonstrated. On the whole it fosters a meritocratic view of power, based on the ability to deliver, rather than being against power as such, and it is this which often makes it an agent of managerialism. This focus on delivery is an immanent aspect of the western capitalist concern with productivity and 'performance' as a dominant value. Going back to Max Weber, evaluation seeks to measure whether the capitalist has earned his or her place among 'the elect'. Thus though traditional evaluation has critical features its narrow focus on immediate performance often makes it part of the structure of contemporary power rather than a radical challenge to it. Nevertheless, there is an element of instability here, in that questioning the value of *some* medical procedures, for example, can lead on to a more general scepticism. Evaluation, for example, asks decision makers to substantiate claims that there are 'weapons of mass destruction' which can be launched within 45 minutes. Evaluation insists that the emperors show their clothes, and this means that its critical potential can never be subsumed within a managerialist imperative. Evaluation does not take things on trust, but asks for the evidence, and this is what makes it potentially a subversive activity.

So if evaluation is often part of the 'end of ideology' or de-politicisation of social life this is a tendency which is not always easily accomplished by those in power, even though they have, among other things, the power to commission evaluations and establish policy problematics. However if we are to make evaluation potentially part of a more thorough-going democratic critique of contemporary power structures, then there is a need to shift the focus on performance away from just asking what things work, how and why. We also need to try to use our human and ethical qualities of judgement to imagine what is possible and what is worth doing. Thus a reflexive approach to

evaluation seeks to connect the evaluation of specific interventions to their 'high level' objectives, and lays these out for potential critical judgement. It seeks to tease out the fact that ideologies and values *always* set the context for evaluation activities 'lower down the line'. If evaluation is seen as a rational and critical activity then it must start in this way. It also needs to ask who sets these objectives, and more especially whether their assumptions are problematic in some way or other. Evaluation always takes place in a power context. Following Foucault (see Rabinow 1986) we can accept that power relations are unstable, and deconstruction of discourses can show how contradictions and inconsistencies create the basis for resistance.

However, what I would also argue is that there is a level of analysis to undertake beyond deconstruction, to overcome tendencies which seem to be manifest within poststructuralism and postmodernism towards relativism and even negativism towards the possibility of policy improvement. These are of course complex issues and there is a danger of setting up 'straw' postmodernists to easily knock them down. Nevertheless a tendency to relativism can occur as a result of the 'linguistic turn' with discourses seen as language constructions that seek to order reality to conform to the power gambits of the powerful groups who mobilise them through 'discursive practices'. This is allied to seeing social practices as 'social constructions' which are politically loaded terms like 'workshy' or 'underclass' whose power lies in the persuasive ability of such categories to shape social understandings and associated means of intervention. There is no independent social reality which is not contaminated by language and power. Of course it is the case that struggles around language and meaning, to impose one set of meanings or another, can have a profound effect on social relations (Fairclough 2001). For example, behind the concept of the workshy lies the 'unquestionable' ethical principle that there is a social responsibility to undertake paid market work, and behind that the legitimacy of the capitalist order itself. Nevertheless from a realist perspective, as well as showing the power that such discourses exert, there are scientific and evaluative questions to answer as best one can, on whether it is the case that unemployment in market societies is a result of an unwillingness to work or not. It could be argued that much unemployment is not caused in such a way or that some of it is but celebrated instead as 'resistance

to wage labour'. These questions or issues of course can only be approached through language constructions, but that does not prevent us from making scientifically validated statements about them that seek to be 'up front' or reflexive about the value principles that also underpin them. In such a way the significance of language and discourse can be acknowledged, without preventing a social scientific understanding of causal processes.

If a tendency towards relativism derives from a suspicion towards the Enlightenment doctrine of scientific truth, a tendency towards negativism derives from a scepticism towards the Enlightenment project of social progress. If professional and state discourses are seen as means of mobilizing the power of some group or agent in relation to vulnerable subjects, they are less likely to be seen as having positive social effects and there is also likelihood that resistance to them will be celebrated as healthy opposition to an oppressive force. Thus in place of seeing 'health' as something definable and positive, poststructuralist thinkers after Foucault (Rabinow 1986) may rather conceptualise it as a questionable political definition mobilised by a powerful group, who are seeking to use it to justify surveillance and even control. The poststructuralist tradition particularly critiques the way that power operates through rational 'liberal' discourses that seek to obtain the collaborative consent of the subjugated through processes that Foucault calls 'governmentality', whose success lies in the individual self regulation of behaviour as a result of discourse, for example, if we seek to avoid the shame of being workshy and take action to be healthy to avoid the stigma of being sick.

I am therefore suggesting that despite many conceptual strengths, there is a danger that in criticizing Enlightenment traditions of unquestioning faith in scientific reason and progress, postmodern and constructivist forms of explanation can have an implicit anti-Enlightenment set of biases that go well beyond scepticism towards rejection. This can get in the way of a balanced assessment that recognizes, on the one hand, that health may have the potential to be an oppressive form of dependency mobilised by doctors, and that the responsibility to work can be a manipulative ethical principle within market capitalism. Nevertheless, on the other hand, doctors can in some circumstances do good, health can be a social benefit, and paid work be a route to material and social betterment. It is the purpose of evaluation therefore to

move beyond both positivist certainty and postmodern relativism to help us to assess these issues.

Thus, in summary, increasingly influential critical traditions view knowledge as inherently uncertain and ‘non-scientific’, and therefore traditional evaluation activities are themselves contaminated by power. It is the critique of knowledge or deconstruction that is empowering, and evaluations which do not acknowledge this may be regarded as technologies of power and domination. Linked to this is the critique of measurement and a preference for qualitative methods. Evaluation is then seen as involving the identification of various forms of evidence and stakeholder perspectives of how projects are operating, without any sense that the evaluator is an expert judge who can decide between them. Such ‘stakeholder’ evaluations may accept that everyone’s reality is equally valid, and that the task of the evaluator is to gain recognition of this from participants, and achieve a pluralism in the way that projects operate. This approach has echoes with the development of ‘collaborative’ approaches to evaluation (O’Sullivan 2004). Linked to this may be an unwillingness to go beyond ‘surface’ explanations to uncover ‘deep’ explanations.

In recent years however, other critical traditions have emerged, which represent both an alternative to both positivist and constructivist approaches. An alternative ‘realist’ or ‘critical realist’ critique of traditional methods however, suggests that it is possible to gather forms of evaluation evidence, both qualitative and quantitative, which is consistent with ‘triangulation’ to uncover ‘deep’ knowledge, mechanisms which may not be immediately apparent to the participants, forming broader contextual influences ‘outside’ the project. These would be missed if we simply sought to reconcile competing stakeholder perspectives or discourses. At the same time project processes, including the views or discourses influencing project agents, need to be seen as factors shaping outcomes. This does however acknowledge the significance of discourses and power frameworks in setting up the way that evaluation questions are asked, and require us to ask in whose interests something is said to ‘work’ or not, as well as of course how findings are used, marginalized or suppressed (see Pawson and Tilley (1997) for a fuller exposition). Nevertheless realist approaches are underpinned by the notion that evaluation



knowledge, though often servicing elites, can make verifiable statements about the world, and in the right circumstances can be utilized to subvert power structures and achieve social transformation. This connects to wider debates in social science and philosophy, notably that developed by Bhaskar (1979) and Layder (1994). While the realist approach is the stance taken within this article, see Burr (2003) for a rigorous exposition and defence of the constructivist perspective.

Evaluation research is inherently applied. Traditionally it embodies the notion that social intervention to produce change is justified and plausible. Evaluation is thus in the first instance an 'interventionist discourse' and like Karl Marx it is not only in favour of philosophical interpretation of the world but seeks to change it. The concept of praxis identifies the relation between intervention and knowledge as dialogical. This could be taken a stage further in that our understanding of the boundaries of what is socially given by structure and made socially possible by agency can only be uncovered in the sphere of intervention. Evaluation can therefore offer both advice to programmes on their effectiveness, and contribute to social science knowledge. While realist inspired evaluations would seek to retain this applied agenda, linked to conceptions of progressive knowledge informing progress in practice, postmodern approaches might tend to be less sanguine, and more likely seek to show that all interventions are in some way uncertain, even flawed, and that the notion of progress an illusion.

Some argue that the purpose of evaluation is not to produce new knowledge:

The most important purpose of evaluation is not to prove but improve (Stufflebeam and Shinkfield 1985: 151).

Evaluation, unlike the basic sciences, does not aim for 'truth' or certainty. Its aim is to help programming and policy making (Weiss 1997:516).

In the UK, such assertions certainly chime in with current 'New Labour' and pragmatic approaches to evaluation focusing simply on 'what works'. While claiming to be non-ideological, they are in fact consistent with traditional empiricist approaches to scientific knowledge and evaluation. They are ideo-

logical in the sense that they take broader structures of power for granted, and only seek to make small changes and adaptations within them.

The critique of empiricist assumptions and a focus on broader rather than narrower issues has led to an emphasis on what is called ‘theory driven’ evaluation. Rather than just trying pragmatically to find out ‘what works’, it also wants to know *why*. In so doing so it seeks to break down the artificial distinction between basic and applied research (Chen/Rossi 1981). Part of this is an assertion of the need for programmes to make *explicit* what is usually left implicit about how their interventions are expected to work, in other words, the *causal* relationships involved, as:

Programme theory is generally made up of a combination of hunches, beliefs, intuitive assumptions and knowledge founded on practical experience (Clarke 1999: 31).

Theory driven evaluation therefore seeks to develop a broader agenda of causes – ‘theories of change’ – involving the interaction of complex contextual and project effects, of which the ‘realistic evaluation’ approach developed by Pawson and Tilley (1997) – considered later in this article – is an extension.

The point that I wish to make about this broader agenda is however political as much as scientific. It is not just that the awareness of complex contextual influences is scientifically superior to the empiricist and pragmatic approaches to evaluation. It is also the case that awareness of broader social and political influences keeps open the possibility that action needs to be taken at a broader level, not just the discrete programme level. In other words, evaluators can point out that contexts themselves are in need of interventionist change. This broader agenda was nowhere better expressed than by C Wright Mills (1959) who – using unemployment and divorce as examples – pointed out that such ‘personal troubles’ if they affect large numbers of people in patterned ways, could be regarded as ‘issues of social structure’. This is also consistent with a ‘realist’ agenda that assumes that the ‘real’ causes may not always be apparent to those involved at the microlevel.

The main arguments I am trying to develop are:

- There is no ‘one best way’ approach to evaluation but different frameworks and paradigms;
- These connect to wider social science, philosophical and political debates;
- In choosing an approach, there is a need to be ‘reflexive’ towards such issues;
- While the traditional models are deeply problematic, systematic and even scientifically founded evaluation is still possible;
- Rather than seeing evaluation as simply applied and atheoretical, it should be ‘theory driven’, and indeed can reflect back on evaluation and social theory itself;
- Analysis of contextual effects have policy and political implications as well as scientific ones.

One final observation is that many of the traditional approaches to evaluation, as well as their critiques, originate in the USA. This is itself interesting, and indicates how the ‘evidence-based’ movement is an aspect of globalisation. As Fairclough (2001: 206) points out globalisation is both ‘process’ and ‘discourse’. In this regard evaluation needs to be treated cautiously as part of a universalising mission associated with neoliberal North American capitalism. In many ways the techniques of evaluation often do represent the extension of neoliberal economic rationales into the spheres of statecraft and governance of social programmes, though I am not suggesting that all evaluation activities can be reduced to this project. To the extent that they are, however, maybe we can develop our own European alternatives linked to preserving and improving the European social model, not rejecting the ‘evidence based’ movement as such but developing one that is not mechanistic and individualistic, but holistic and social. Above all one that puts evaluation tools into the hands of people themselves rather than new cadres of administrators. I will come back to this later when I discuss specific frameworks for looking at employment initiatives.

While evaluation can be seen as one of the concomitants of the increasing rationalisation associated with capitalist modernity, recent trends have inten-

sified its application. These include the imperatives to rationing that have arisen as a result of the 'fiscal crisis' of the welfare state (O'Connor 1973), that the demands of social programmes arising from capitalism, both the problems it generates, and the expectations that it raises politically, are more than tax systems can fund. Therefore means have to be found to decide which should be prioritised, in a legitimating as well as a technical sense. However, in case it is thought that I am being economically reductionist, I do acknowledge that evaluation pressures also arise from the critical activities of modernity itself. For example, I have already referred to Cochrane's (1972) classic critique of medicine's professional interventions that they were often based more on hunches than scientific testing, which advocated using the 'gold standard' scientific method of experiments or 'randomised controlled trials' to assess which are useful, irrelevant or damaging – about which more below. This new sense of uncertainty and lack of trust could be seen as an aspect of an enhanced reflexivity to risk associated with what Giddens (1991) calls 'late modernity' and Beck (1992) 'risk society'.

These are therefore some of the wider material and ideological influences on the 'evidence-based movement'. If realist evaluation is about taking the broader context into account, it undoubtedly also needs to interrogate more deeply the reasons why evaluation imperatives have become intense. In other words, the history of evaluation is context bound, and the shifts in evaluation paradigms are shaped by wider ideological and political-economic changes.

### **3. Four generations of evaluation research?**

Thus rather than searching for the 'one best way' to undertake evaluation research, a reflexive model emphasises the development of different 'paradigms' of evaluation, and also seeks to analyse the contexts in which they emerge and are contested.

Within a 'political' approach some paradigms may be seen as enhancing possibilities for control from below rather than control from above. The most useful starting point in this regard is the Guba and Lincoln (1989) evolutionary typology of four generations of evaluation research which can be said to have helped to politicize evaluation from a constructivist perspective. They

are not saying that one generation necessarily overthrows the next, as clearly the first three are still prominent and indeed influential in evaluation circles. They interestingly show that one of the mainsprings of evaluation methods, which has now spread tentacles into many other areas, often came initially from efforts to provide an evidence base for educational policies, from which most of their examples are drawn.

### ***The first generation: Measurement***

Guba and Lincoln show how this emphasis on measurement came out of particular trends which gathered pace at the turn of the 20<sup>th</sup> Century, particularly the measurement of schoolchildren's aptitude which started in Britain and France and spread to the USA. Also the drafting of large numbers of soldiers for World War I in the USA led to the spread of intelligence and personality testing of adults. Alongside this the scientific management movement in industry sought to assess and measure the productivity of workers. Within this system they argue that evaluators had primarily a technical role to refine and administer the battery of tests. Measurement and evaluation became seen as synonymous. The USA has particularly led the way in developing such standards for example since the first *Stanford Achievement Battery* in 1922. In other spheres, I would suggest that US capitalism has put special emphasis on classification in many spheres, for example in mental health it has led to the voluminous *Diagnostic and Statistical Manual (DSM)*, whose various editions have sought to define the professional parameters of the full range of mental disorders. As these become international standards, they become an aspect of globalisation.

I would also claim that the context in which educational evaluation sought to measure educational performance came out of efforts to define 'educability', and which children from lower social classes it was worth investing public money and effort upon. Behind this often lay ideas that intelligence was primarily innate, which was also frequently embedded in eugenicist legitimations of an unequal national and international social order, i.e. often racist as well as class biased. This focus on measurement could be framed in poststructuralist terms as part of a development of range of policy and profes-

sional activities which Foucault, in a telling phrase, calls ‘dividing practices’ (Rabinow 1986).

### ***Second generation: Description***

Guba and Lincoln also show that the educational context in the USA was significant in facilitating a shift of the evaluator into a more proactive ‘descriptive’ role, rather than simply being someone who passively engaged in measurement. This involved a shift of focus from individual pupils and students, towards *programmes*. Thus Ralph W. Tyler of Ohio State University is seen as the ‘father of evaluation’ who developed tests to see whether a curriculum itself enabled students to learn or not, that is, asked whether programmes achieved what they set out to do or not, and whether they therefore need modifying. This came out of the Eight Year Study of secondary schools initiated in 1933. What he developed later became known as ‘formative evaluation’, that is analysis of strengths and weaknesses of programmes in relation to objectives some time after they had been in operation. It established evaluation as a systematic process starting from objectives, testing them against outcomes and then fine tuning the programme in the light of them (‘Tyler’s rationale’). His work also started to draw attention to the social and organisational effects on programme outcomes, and therefore can be seen as the beginnings of systematic process evaluation. However he operated with a firm notion that programmes were underpinned by consensual objectives.

### ***Third generation: Judgement***

Guba and Lincoln argue that concerns at the performance of the US educational system in the context of the Cold War, specifically the fact Russians got someone in space first, led to a state of moral panic which enhanced the role of the evaluator to make ‘judgements’ about the objectives before programmes started, rather than leaving it until it was too late to make difference. A key landmark again in the educational field was Cronbach’s (1963) classic ‘course improvement through evaluation’. I would also argue that the growing fiscal pressure on social programmes in this era as identified by O’Connor (1973) and the allied beginnings of a challenge to professionalism

by managerialism, were also influential. The concern with performance was indeed operating in a global context of competition between rival social systems, as well as at a domestic level. It can be seen as the signs of an awareness that simply providing access to programmes was not necessarily sufficient. The social democratic consensus of the post 1945 world had often been underpinned by an implicit faith in the effectiveness of programmes, and in the 'experts' who administered them.

***Fourth generation: 'Responsive constructivist evaluation'***

Guba and Lincoln argue that all three approaches are primarily 'managerialist' in that they assume a consensus on purposes even if there is some disagreement on methods and the role of the evaluator. Their approach opens up a new and sophisticated evaluation terrain, emphasizing the need to analyse different stakeholders, the 'agents' involved, the direct and indirect 'beneficiaries' and also the fact that some people may be 'victims'. For example, if programmes improve educational standards for some it may be at the expense of others (Guba and Lincoln 1989: 40-1). Their work is broadly consistent with a postmodern shift towards social constructivist relativism and also the position that there is no truth as such to determine beyond the meanings that private actors ascribe to the world (i.e. the 'linguistic turn'). The following quote gives a flavour:

"Evaluation outcomes are not descriptions of the 'way things really are' or 'really work' or of some 'true' state of affairs, but instead represent meaningful constructions that individual actors or groups of actors form to 'make sense' of the situations in which they find themselves. The findings are not 'facts' in some ultimate sense but are, instead, literally *created* through an interactive process that *includes* the evaluator (so much for objectivity!) as well as the many stakeholders that are put at some risk by the evaluation. What emerges from this process is one or more *constructions* that *are* the realities of the case" (Guba and Lincoln 1989: 8).

The key point that they make is that deciding what reality and truth counts in situations where there are multiple viewpoints is not a 'scientific' but a 'negotiated' process. Truth is something which emerges out of the considerations of multiple viewpoints, as facilitated by the evaluators, thus:

“One of the major tasks of the evaluator is to conduct the evaluation in such a way that each group must confront and deal with the constructions of all the others, a process we shall refer to as a hermeneutic dialectic. In that process some, perhaps many, of the original claims, concerns, and issues may be settled without recourse to new information, that is, information that is not already available from one or more of the stakeholding groups themselves. As each group copes with the constructions posed by others, their own constructions alter by virtue of becoming better informed and more sophisticated. Ideally, responsive evaluation seeks to reach consensus on all claims, concerns, and issues at this point, but that is rarely if ever possible. Conflicts remain whose resolution require the introduction of outside information, which it becomes the evaluator’s task to obtain” (Guba and Lincoln, 1989: 41).

Thus it is commitment to a democratic process of arriving at shared meanings through listening to all versions of the truth that is the kernel of the evaluation process.

### ***Realistic evaluation as an alternative to fourth generation evaluation***

The ‘realistic evaluation’ approach developed by Pawson and Tilley (1997) draws heavily on the critique of traditional methods of evaluation, but still asserts a ‘modernist’ focus on causes, rather than simply seeking to arrive at a process of consensus or agreement between stakeholders. They identify four main perspectives on evaluation: experimental, pragmatic, naturalistic and pluralist.

The experimental or ‘foundationalist’ method seeks to take two broadly matched groups, expose one of them to a policy ‘treatment’ and then see what difference has been made by measuring the difference between them. This approach was advocated in relation to social programmes in the USA by Campbell in 1969 through a ‘reforms as experiments’ approach. While the experimental approach may ‘work’ in some circumstances outside the laboratory where the aims are very specific, eg, testing particular drugs, there are severe limitations in testing of complex social interventions where variables are too numerous to control. This led to a shift from a causal to pragmatic and naturalist models as systematic approaches, again developed in the US for research into educational evaluation by Stufflebeam’s 1980 ‘standards of util-



ity, feasibility, propriety and accuracy', that is, utilising a range of appropriate methods in a proper and expert fashion. The pluralist approach is synonymous with 'fourth generation' evaluation where the evaluator moves from being a 'judge' to a 'ring-master', that is someone who reflects back to the stakeholders their own statements in non-judgemental ways, and thereby helps them to understand each others' viewpoints.

Realistic evaluation accepts much of the critique of traditional evaluation while at the same time eschewing the tendency towards relativism, and according the evaluator some role to make judgements based on the principle of 'generative causation'. While their arguments are complex they helpfully summarize in chapter 9 of their book their 'new rules of realistic evaluation' which I will try to distill even further:

- *Need to show both how and why*, and the fact that this is 'triggered' by the interaction of participants in programmes, both providers and recipients. This is an active conception of cause that sees programmes as more than machines, and involving dynamic sets of social relations. Changes in social relations are seen as the chief outcomes of programmes.
- *Generative causes are not always directly observable* – thus stakeholders may not be aware of deeper constraints on their actions, and may need to have these pointed out to them.
- *Need to analyse the effect 'mechanisms' of programmes have on causal mechanisms* – i.e. change itself does not mean that the programme itself is responsible for them. Thus people who receive help to get a job, may have got one anyway. Traditional approaches call this the problem of 'dead-weight' and seek to control for it, whereas realist approaches seek to identify interactive effects.
- *Contextual analysis is vital* – including deciding 'for whom and in what circumstances' a programme's mechanisms work
- *A theory testing approach based on identifying programme 'outcomes'* – this is consistent with identifying the programme's initial 'theory of change' and what events or indicators would serve as criteria of success.

- *Analysis of relations between contexts, mechanisms and outcomes (CMOs) may allow 'lessons' to be learnt and transferred* – the point however is that this should be done carefully
- *Both stakeholders and evaluators can learn from each other* – rather than seeing one or the other as 'the expert'
- *The contextual world in which evaluation takes place is changing all the time* – therefore the notion that it be 'controlled' while evaluation 'treatments' are tested is problematic
- *Evaluation is a craft rather than a 'science'* – thus evaluation knowledge and practice is uncertain, but the evaluation is still a skilled business that offers useful knowledge.

### **A postmodern alternative to realism?**

Those who wish to compare and contrast this realist with a more postmodern approach to evaluation, developed from 'fourth generation' perspectives, could focus on the approach taken by Everitt and Hardiker (1996). As they put it:

“A critical evaluation, then, informed by postmodernist theorizing, suspends the process of seeking causal relationships between inputs and outputs, becoming more aware of ‘the polymorphous cluster of correlations’ that constitute discourses and discursive fields, ‘regimes of truth’ ([citing Foucault – see Rabinow 1986])...The task of evaluation is to contribute to the deconstruction of discourses that serve consistently to render some (for example, young women who are mothers; Black youth) less powerful than others” (Everitt and Hardiker 1996: 106-107).

They recognise that this leads to a charge of relativism, which they answer by advocating a politics of 'needs talk' which draws on the work of Nancy Fraser (1989). This sees needs as neither objective nor subjective but constructed through discourses which can be imposed from above but also generated as 'oppositional needs discourses' from below. In order to decide between these various conceptions of needs Fraser proposes both *procedural* processes which are 'inclusive', and *consequential* approaches which lead to desirable social outcomes such as greater democracy and equality. The first would be

consistent with fourth generation evaluation, though the second goes beyond it. This still however defines needs as primarily 'political' in contrast to the 'ontological' approach developed in the work of Doyal and Gough (1991) in which human need is defined in terms of 'objective' criteria such as health and self realisation in a social context, a point to which I will return in the conclusion.

#### **4. Contextualising and deconstructing EQUAL<sup>1</sup> discourses – Sketching out a realist approach to evaluation**

The argument above compresses complex debates, but I will now try to apply them in outline to the particular case of the European Union's EQUAL Programme. The most significant feature of 'realist' evaluation is a holistic, systems orientation emphasizing: (1) the need to take account of broader political, economic and ideological contexts as causal influences affecting project or programme interventions; (2) the interactive effects between project or programme processes and outcomes, seeing these as dynamic sets of social relationships between participants, rather than mechanical 'treatments'. As far as the first is concerned I seek to utilise a policy analysis framework that looks at the interaction of macro (global or transnational), meso (national) and local (micro) influences, to examine empirically the ways in which they constrain or empower actors involved in programmes. Although this acknowledges the role of the economy as a 'material force', it also takes account of 'new institutionalist' approaches that recognises that markets and their regulatory dimensions are socially made and sustained (for exposition and critique of 'new institutionalist' approaches see Gorges 2001).

At the *macro level* there is a need to focus on the European social policy project and its possibilities and constraints. This emerges out of a global context in which European capitalist elites have since the 1980s felt increasingly challenged by Japanese and US models. In recent years it is particularly the strength of US capitalism, and more belatedly the British variant of it, that has led European elites to believe that European systems are too 'rigid' and

---

<sup>1</sup> For brief outlines of the EQUAL programme and its socio-political contexts see pp. 330 and 332 below.

‘inflexible’ (among others being urged by Gordon Brown, as British Chancellor of the Exchequer, to do so (Seager 2005)). More recently still, the ‘threat’ of Chinese competition has loomed. Europe’s unemployment problem is undoubtedly a ‘real’ and not just a discursive problem, particularly in the three major continental economies in the Eurozone, Germany, France and Spain where International Labour Organisation (ILO) rates were over 9 per cent in August 2005, against an average of 8.6 per cent (Eurostat 2005). This problem, contrasted with low unemployment in the UK (4.6 per cent in June 2005) and the USA (4.9 per cent in August 2005) has been interpreted by orthodox policy analysts as due to the strong influence of unions and over-regulated labour markets, and high costs of social protection, that has not generated the (low wage) dynamic service economy that led to the US boom, or the competitive advantage that China has through low labour costs and authoritarian rule. However while ‘neoliberal’ forces, and political elites influenced by them, would clearly like to see these ‘inhibiting’ structures diminished or dismantled, there is still considerable support among European citizens, trade unions and NGOs for a social model of capitalism, and pressure has in fact mounted to more concerted action by new social movements to combat a range of discriminations and causes of social exclusion. It has been argued that the EU policies are informed by a naive notion that these two conflicting elements – the social and the neoliberal – can be reconciled (see Kleinman, 2002). However, while such scepticism may be justified, this is in fact a key evaluation question in need of an answer.

In analysing this key question in realist ways as a social and not just a mechanical economic process, it needs to be recognized that European decision makers are themselves diverse, involve ‘social partners’ of political and administrative cadres, as well as business interests, and ‘subordinated’ groups such as workers and social movements. Policies are likely to represent a messy compromise between them, and they are unlikely to lead to a ‘pluralist’ levelling out, as some interests, particularly those of multinational corporations, are likely to be ‘more equal than others’ and exert considerable power in shaping agendas in their favour in terms of Lukes’s (1974) ‘radical’ view of power outlined earlier. This is because the currently ascendant discourses, at least at the formal level of European and national political institu-

tions, emphasize above all the need to improve competitiveness and stimulate economic growth. In this context the European Social Model continually has to defend itself against the neoliberal critique, rather than the other way around. The ascendant discourses therefore enhance the power of the most powerful actors and often seem to set the limits within which contestation from below can operate. A realist approach to power therefore posits an interactive relationship between discursive and economic power, in contrast to the one-sided postmodern emphasis on discourse.

The EU's EQUAL Programme emerged out of the Luxembourg Jobs Summit held in November 1997, which launched the European Employment Strategy (EES) which sought to reconcile objectives such as achieving sustained economic growth, full employment, more and better jobs, greater productivity, environmental sustainability and strengthened social cohesion. However the chief means by which these are to be achieved are by raising the employment rate and labour market participation of social groups. At the same time following the Amsterdam Treaty's introduction of the fight against social exclusion among the EU's objectives, the Lisbon European Council of March 2000 asked national governments and the European Commission to take steps to have a significant impact on eradicating poverty by 2010. Both the European Employment Strategy (EES) and the social inclusion strategy have set European wide objectives and requirements to establish National Action Plans. However the means by which these are achieved, through target setting and monitoring, are largely voluntaristic, which in Euro-speak is called 'The Open Method of Coordination' or 'soft law'. This involves setting common objectives and associated indicators, assessment of the impact of national programmes, and a Community Action Programme to promote policy cooperation and transnational exchange of learning and good practice. This article, as stated in the introduction, is one micro-product of this process! At the larger scale, it has been acknowledged that many of the Lisbon targets which the EU set itself have at best been only partially met at European or national level. In 2005 the EU therefore decided to relaunch the Lisbon strategy as it has been recognized that there have been problems of 'implementation' (Euractiv 2005).

This would seem to imply that the problems are primarily technical ones rather than open to political interpretation. In answering the key evaluation question about the extent to which a neoliberal and social Europe can be reconciled, there is a need to assess why the Lisbon strategy has been at best only partially successful against the criteria it set itself, including raising employment and lowering unemployment, and also why it has been more successful in some places than others. All I will do here in the limited space available is to map out two possible emerging explanations, which criticise the strategy for being too timid in one or another direction. The first neoliberal explanation is that there is a need for a more thoroughgoing liberalisation and removal of barriers, for example, Gordon Brown as cited above. The second explanation is neo-Keynesian arguing that the public policy instruments used to reconcile social objectives with economic adaptation are not bold or robust enough. In particular the Open Method of Coordination is seen as leaving too much to national governments while at the same time dominant 'stability' focused EU level instruments aimed at keeping inflation low have remained largely intact. This latter view tends to be promoted by leftist and 'subordinated' social actors such as the European Trade Union Council (ETUC). What this illustrates of course at the macro level is the political nature of any explanation and the need for any evaluation of the Lisbon strategy to be reflexive about ideological or explanatory starting points. There is also a need to produce a differentiated rather than 'British versus continental Europe' account. While it is the case that Britain since the mid 1990s has combined flexible labour market reforms with low unemployment, Scandinavian countries like Sweden, (6.3 per cent in March 2005) and particularly Denmark (4.8 per cent in August 2005) have also done reasonably well without radically dismantling their welfare systems (Eurostat 2005). According to some, lefts greens etc, the relaunch of the Lisbon strategy is prioritising growth over social objectives (Euractive 2005).

Within this complex macro-interventionist framework, EQUAL is just one relatively small element linked to the European Employment Strategy (EES) through the European Social Fund (ESF). The EQUAL Programme is defined as a 'laboratory of new ideas' which will then lead to learning and implementation by both the European Employment Strategy (EES) and the Social In-

clusion Process. It seeks to do this through the pragmatic approach to evaluation described earlier in this article. In other words it seeks to pioneer relatively small scale but promising practical approaches to combat discriminations and promote 'employability', 'entrepreneurship', 'adaptability' and 'equal opportunities'. The aim is then to demonstrate 'what works' at project level and across transnational networks, which will then be 'mainstreamed' by policy makers. While the focus is about achieving social inclusion through participation in the labour market, EQUAL broadens the agenda by focusing on efforts to combat forms of discrimination and exclusion based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation (see EQUAL homepage: <http://europa.eu.int/index.eu.htm>).

These overall objectives of the European Employment Strategy (EES) and the Social Inclusion Process, as well as the specific features of the EQUAL Programme, involve efforts to create a consensus among the participants to create an agreed system of centralised monitoring and intelligence gathering. In this respect they seem to fall at central level into a large scale form of what Guba and Lincoln would categorise as 'third generation' or 'judgement' focused evaluation. Though there are elements of 'responsive constructivist evaluation' through involvement of social partners and the primarily voluntaristic nature of the Open Method of Coordination, there is a nevertheless a set of common assumptions that underpin the whole process which are not seen as open to evaluation or contestation. There is no admitted contradiction between the European Employment Strategy (EES) and the Social Inclusion Process, not only because the Lisbon strategy clearly states that 'employment is the best safeguard against social exclusion', but also because supply side policies are seen as the prime means of achieving it:

In order to promote quality employment it is necessary to develop employability, in particular through policies to promote the acquisition of skills and life-long learning (European Council, 2000).

It is also important to engage in *meso or national level* analysis, as it is widely recognized that, despite the moves to European integration, there are significant differences between 'policy regimes' at national level in Europe, and the Open Method of Coordination gives scope for different national interpretations. As Kleinman (2002) asks, is there 'one social model or many'?

in which case Britain's model now looks among the least social in Europe, closer to a North American 'liberal' approach in placing an increasing primacy on compulsory participation in the labour market. Thus in Britain, which has only experienced short periods of 'weak' European style corporatism, and much of which was in any case stripped away in the Thatcher era, neoliberal labour markets are a prominent feature. These can be characterized by US style features of deregulation, growth of the service economy and low wage employment, and a dynamic economy and tight labour markets. Nevertheless British forms of social protection and employment assistance through policies such as New Labour's New Deal for Employment still mark it as distinct from the USA. This is not to suggest, however, that the UK's approach is fundamentally at variance from the European wide strategy which as we saw particularly emphasizes employability and supply side strategies as ways of tackling poverty and exclusion. This then forms the key features of the context in which the European Employment Strategy (EES) in general and EQUAL programmes in particular operate in the UK.

Thus EQUAL programmes may incorporate contradictory objectives seeking to reconcile social objectives with the free market to varying degrees and it will be significant to examine the extent to which local projects in different settings are constrained by these or not. There is certainly a lot of 'needs talk' going on around them which it would be useful to deconstruct, with some discourses more 'top-down' and others 'bottom up'. The key evaluative question in this regard is whether integration into labour markets is the key issue to address in terms of combating discrimination and exclusion. People may be empowered to take jobs but may also be disempowered by being forced to do so by restrictive benefit regimes which give them no other choice. For some groups in such circumstances, e.g. forced integration of single parents into the labour market may represent a loss of power to balance paid work against parenting concerns, and this may heighten gender discrimination. The incentives to take jobs are also affected by their quality, and in Britain there has been a relative decline in the quality of jobs available at the lower end of the market in terms of wages level, security, and discretion on the job (Gregg and Wadsworth, 1999). The key issue here from a more



‘third generation’ evaluation perspective, would be to see what people’s own agendas and aspirations are, and whether they have ‘realistic’ perceptions of costs and benefits of employment. In other words, how they define their own ‘needs’ in relation to employment and life goals.

In terms of an evaluative model, in Britain official agencies like Job Centre Plus see entry into jobs as the key criteria of success, whereas the Learning and Skills Council defines it a little more flexibly in terms of participation in training schemes. More recently alternative models such as ‘distance travelled’ have been developed to recognize that progress towards integration into the labour market may be a staged process and need to be measured by ‘soft outcomes’ such as time-keeping, communication skills, improved esteem etc as well as the ‘hard outcomes’ such as finding a job (Lloyd and O’Sullivan, 2003). Such measures it could be argued are more favourable to pluralist and realist evaluation models in that:

- They recognize that causes of labour market exclusion are complex and that some ‘intermediate’ outcomes may contribute to eventual job success;
- They embody ‘theories of change’ themselves about the reasons for unemployment, that represent a degree of sophistication;
- They open up the possibility that for some people the goal should not necessarily be employment as such, and that a humanistic and holistic conception of self development is a more important aim than just getting a job.

Set against these possibilities, however:

- The division between ‘hard’ and ‘soft’ outcomes places a hierarchy on the first, so that getting a job is seen as the ultimate ‘gold standard’ aim;
- There is still a fetish of measurement even with regard to soft outcomes which mitigates against a holistic approach;
- The ‘theory of change’ itself takes little account of *contexts* but assumes progress will occur towards the desired aims if the people themselves are

enabled by projects to adopt the ‘right’ characteristics and to travel along the road laid out for them in the approved way

- These desired aims are not *pluralist* in the sense that notions of progress and distance travelled are those imposed by the evaluators, rather than arising from the agendas of participants, and analysis of the journeys they might want to devise for themselves.

## 5. Conclusion

The previous section has hopefully sketched out the beginnings of a realist evaluation approach to EQUAL Employability programmes, which could be applied at European-wide, national and local or project levels, in ways that utilise the insights provided by constructivist or discursive approaches. The aim has been to show the value of deconstructing the discourses underpinning social programmes, and also the evaluation ‘paradigms’ that are linked to them. This could enable a more thorough-going critical evaluation *of* programmes, as well as evaluation *within* them. In terms of the European Employment Strategy (EES), it would lead to a focus on identifying the potential contradictions as well as the harmonies between pursuing poverty and exclusion in tandem with raising the employment level, as well as the limits of supply side approaches focusing primarily on improving the employability of labour market participants. Such evaluative analyses can help facilitate a wider conception of political choices, informed by a more sophisticated notion of the links between poverty and exclusion and involvement in the capitalist labour market, than seem to be ‘permitted’ by current EU policy discourses. I would of course hope the evaluation approach advocated and developed here might be mainstreamed as an example showing the benefits of starting by deconstructing the discourses of social policies, programmes and projects, but then going on to evaluate the causes of success or failure through appropriate and reflexive social scientific methods.

What this could do is open up a productive debate and discussion between ‘critical’ evaluation approaches, where much of the force of ‘fourth generation’ and postmodern critiques is acknowledged, but the possibility of ‘realism’ is retained, and along with it the notion that the evaluator may independently

have something of significance to uncover, and that evaluation itself may contribute to restoring the credibility of social intervention as a contributor to social progress and emancipation. As part of this, there needs to be both extensive broad contextual analysis, efforts to uncover the voices of those who are targeted by social programmes, and also – here’s the challenging bit – laying out means by which these two elements can be brought together within a causal framework. Stakeholder approaches with their emphasis on qualitative methods undoubtedly make this possible, and with other colleagues, I tried to show in the context of unemployment initiatives in Coventry, UK, how the voices of project workers and clients can be used to draw out policy messages for mainstreaming (Aleksandraviciene et al. 2005).

A final issue, in the context of this journal, is the linkage of all this to action research. The strong message coming from this article is the need for inclusive stakeholder methods, for project actors to be involved in defining their ‘theories of change’ from the outset and for them to be reflexive about their interests and ideological starting points. Subsequently they need to continuously review progress against agreed outcomes, combined with awareness of how these are affected by wider contexts and organisational processes, facilitating a reflexive learning strategy. This makes realist and collaborative evaluation a social activity which aims to maximise the chances of project success rather than just judge outcomes from a distance, and therefore makes it an integral part of project praxis, ideally generating cycles of interdependence between theory, evidence and intervention.

## References

- Aleksandraviciene, I. / Cleaver, P. / Hussain, Z. / Jovanovic, J. / Merrill, B. (2005): More than Just a Number: The Impact of Community Based Initiatives on Unemployed Men and Women in Coventry, Final Report of SEQUAL Project on Gender and Social Class. Accessed at: [www.surrey.ac.uk/politics/cse/sequal.htm](http://www.surrey.ac.uk/politics/cse/sequal.htm).
- Bhaskar, R. (1979): *The Possibility of Naturalism*. Brighton: Harvester.
- Burr, V. (2003): *Social Constructionism*. London and New York: Routledge.
- Chen, H. T. / Rossi, P. H. (1981): The Multi-Goal, Theory-Driven Approach to Evaluation: A Model Linking Basic and Applied Social Science. In: H. E Freeman and M. A. Solomon (eds.): *Evaluation Studies Annual*. Vol 6. Newbury Park, CA: Sage.

- Chen, H. T. / Rossi, P. H. (1983): Evaluating with Sense: The Theory-Driven Approach. In: *Evaluation Review*, 7: 283-302.
- Chen, H. T. (1990): *Theory Driven Evaluations*. Newbury Park, CA: Sage.
- Clarke, A. / Dawson, R. (1999): *Evaluation Research: An Introduction to Principles, Methods and Practice*. London: Sage.
- Cochrane, A. (1972): *Effectiveness and Efficiency: Random Reflections of the Heralth Service*. London: Nuffield Provincial Hospital Trust.
- Coote, Anna / Allen, Jessica / Woodhead, David (2004): *Finding Out What Works: Understanding Complex, Community Based Initiatives*. London: King's Fund.
- Cronbach, L. (1963): Course improvement through evaluation. In: *Teachers College Record*. 64, 672-683.
- Davies, H. T. O. / Nutley, Sandra M. / Smith, Peter C. (2000): *What Works? Evidence-based Policy and Practice in Public Services*. Bristol: Policy Press.
- Doyal, L. / Gough, I. (1991): *A Theory of Human Need*. London: Macmillan.
- Euractive (2005): *Relaunch of Lisbon Strategy*. Accessed at <http://www.euractiv.com/>.
- Eurostat (2005): *Euro-zone Unemployment up to 8.6 Per Cent*. Euro-indicators News Release, 4 October. Luxembourg: Office for Official Publications of the European Communities. Accessed at <http://epp.eurostat.cec.eu.int/>.
- European Council (2000): *Fight Against Poverty and Social Exclusion – Definition of Appropriate Objectives*. Luxembourg: Office for Official Publications of the European Communities. Accessed at <http://epp.eurostat.cec.eu.int/>.
- Everitt, A. / Hardiker, P. (1996): *Evaluating for Good Practice*. London: Palgrave.
- Fairclough, N. (2001): *Language and Power*. Harlow: Pearson Education, second edition.
- Fraser, N. (1989): Talking About Needs: Interpretative Contests as Political Conflicts in Welfare State Societies. In: *Ethics*, 99 (January): 291-313.
- Giddens, A. (1991): *Modernity and Self-Identity: Self and Society in the Late Modern Age*. Cambridge: Polity Press.
- Gorges, Michael J. (2001): New Institutional Explanations for Institutional Change: A Note of Caution. In: *Politics* 21(2): 137-145.
- Gregg, P. / Wadsworth, J. (1999): *The State of Working Britain*. Manchester: Manchester University Press.
- Kleinman, M. (2002): *A European Welfare State? European Union Social Policy in Context*. London: Palgrave.
- Layder, D. (1994): *Understanding Social Theory*, London: Sage.
- Lloyd, R. / O'Sullivan, F. (2003): *Measuring Soft Outcomes and Distance Travelled: A Practical Guide and Existing Models*. London: Department for Work and Pensions.
- Lukes, S. (1974): *Power: A Radical View*. London: Macmillan.

- Mills, C. Wright (1959): *The Sociological Imagination*. New York: Oxford University Press.
- O'Connor J. (1973): *The Fiscal Crisis of the State*. New York: St James's Press.
- O'Sullivan, Rita G. (2004): *Practicing Evaluation: A Collaborative Approach*. Thousand Oaks: Sage.
- Pawson, R. / Tilly, N. (1997): *Realistic Evaluation*. London: Sage.
- Rabinow, P. (ed.) (1986): *The Foucault Reader*. Harmondsworth: Penguin.
- Stufflebeam, D. L. / Shinkfield, A. J. (1985): *Systematic Evaluation: A Self-Instructional Guide to Theory and Practice*. Dordrecht: Kluwer Nijhoff.
- Taylor, D. /d Balloch, S. (2005): *The Politics of Evaluation*. Bristol: Policy Press.
- Weiss, C. (1997): How Can Theory-based Evaluation Make Greater Headway? In: *Evaluation Review*. 21(4): 504-24.

*About the author*

Mick Carpenter is Reader in Social Policy in the Department of Sociology at the University of Warwick. He has researched the effects of community based initiatives in combating unemployment in Coventry through the SEQUAL project, and other initiatives evaluating social change in the city, through his work for the Coventry Partnership Community Research and Evaluation Service (CRES).

*Author's address*

Mick Carpenter  
 Department of Sociology  
 University of Warwick  
 Coventry CV4 7AL  
 UK  
 E-mail: m.j.carpenter@warwick.ac.uk